模糊数学评价河流水质的模型探讨

孙 蕾

（湖南省环境监测中心站，长沙 410004）

摘 要 采用模糊数学中模糊聚类的2种方法即最大矩阵元法和模糊相关分析法对武水河水质进行分析、聚类和评价。试图通过实例阐明最大矩阵元法和模糊相关分析法对河流水体单元、尤其是污染参数进行聚类分析评价，从不同侧面全面剖析河流水质污染状况，分析污染因子在河水中的相互作用及各种物理化学作用所形成的污染参数之间的相互关系。

关键词 水质 模糊聚类分析 评价

水质评价是环境科学中广泛使用的一种表述水体水质状况的手段，其评价方法目前通用的是单项污染指数法。单项指数法作为一种标准方法具有简明、操作性强等特点。但为了更深入全面地分析河流水质，不妨在水质评价方法学上进行更多的探讨。

水体污染程度和污染因子在河流中的相互关系等都客观存在着模糊性。用模糊数学中模糊聚类的2种方法从不同角度对水体水质进行剖析，对水体单元和污染因子进行聚类分析，试图通过这2个模型，建立用模糊聚类分析对水体单元和污染物进行聚类评价的通用方法。

1 模糊数学进行水质评价的理论和模型

1.1 模糊矩阵及其建立

定义矩阵 $R=(r_{ij})_{n\times m}$ 叫做一个模糊矩阵，对任意 $i=1, 2, \cdots n$，$j=1, 2, \cdots m$，都有 $r_{ij}\in[0, 1]$。

用模糊数学进行水质评价，建立模糊矩阵是一个核心工作，对于任何一个研究水体，都有两个因素构成，即水体单元和各水体单元中的污染参数。其模糊相关矩阵的矩阵元 r_{ij} 的确定如下：

设 $X=(x_1, x_2, \cdots x_n)$ 为待分类事物的全体，即水体单元，x_1 由一组数据 U_{i1}，U_{i2}，$\cdots U_{im}$ 表征。X 上的模糊相关矩阵 $R=(r_{ij})_{n\times m}$ 可用多种方法计算。在此，为建立两个水体单元之间的相似关系，采用模糊集中的“贴近度”概念之一的最大最小法计算：

$$r_{ij}=\frac{\sum_{k=1}^{m}\min(U_{ik}, U_{jk})}{\sum_{k=1}^{m}\max(U_{ik}, U_{jk})}\quad (i, j\leqslant n)\qquad(1)$$

上式中水质因子对目标值隶属度 U_{ij} 的确定按下列公式计算：

$$U_{ij}=\begin{cases}\dfrac{S_{ij}}{C_{ij}} & C_{ij}>S_{ij}\\ 1 & \text{实测值优于目标值时}\end{cases}$$

式中：U_{ij}——i 水体单元 j 类污染物对目标值的隶属程度；

C_{ij}——i 水体单元 j 类污染物的实测浓度；

S_{ij}——i 水体单元 j 类污染物的目标值。

1.2 污染因子的模糊相关分析

对于环境问题，在引起环境变化的诸因素中，往往有未知的或无法监测的，就是在已监测的因子中，很可能有些相互没有实质性的联系或由于某种干扰使得有实质性联系的因素未显示出来。因此有必要采取一定的数学方法进行筛选。

以相似系数法建立模糊相关矩阵。由此矩阵可分析各样本间的相互关系，即各污染因子间的相关性。但应注意的是，此处的相关是两个污染因子之间在量上的相互依赖程度，至于这种依赖是正相关还是负相关则需要依其物理化学机制而定。

在已建的目标值隶属度的基础上，以相似系数法建立模糊相关矩阵，其矩阵元 r_{ij} 为：

$$r_{ij}=\frac{\sum_{k=1}^{m}(U_{jk}-\overline{U}_j)(U_{jk}-\overline{U}_j)}{\sqrt{\sum_{k=1}^{m}(U_{ik}-\overline{U}_i)^2}\cdot\sqrt{\sum_{k=1}^{m}(U_{jk}-\overline{U}_j)}} \quad (2)$$

式中：$\overline{U}_i=\frac{1}{m}\sum_{k=1}^{m}U_{ik}$，$\overline{U}_j=\frac{1}{m}\sum_{k=1}^{m}U_{jk}$

2 应用实例

2.1 最大矩阵元法对武水河水质的评价

2.1.1 水体单元的聚类评价

以监测断面为水体单元，根据各水体单元水质监测值建立各水体单元的水质因子对环境目标的隶属度，由隶属度求得的模糊相关矩阵为：

$$R=\begin{bmatrix} 1 & & & & & & \\ 0.96 & 1 & & & & & \\ 0.98 & 0.94 & 1 & & & & \\ 0.99 & 0.96 & 0.98 & 1 & & & \\ 0.98 & 0.95 & 0.98 & 0.98 & 1 & & \\ 0.94 & 0.91 & 0.93 & 0.94 & 0.94 & 1 & \\ 0.95 & 0.93 & 0.95 & 0.95 & 0.95 & 0.93 & 1 \end{bmatrix} \quad (3)$$

以上矩阵为模糊相似矩阵，模糊相似矩阵具有自反性和对称性，但不具有传递性，只有将模糊相似关系通过求其传递闭包而转换为模糊等价关系，矩阵的任一截矩阵 R_λ 才具有传递性，即有 R_λ 为 λ — 截矩阵（布尔矩阵）。经过 R 矩阵的变换，得 R 的等价矩阵如下：

其中 λ 为置信水平。对于一个等价矩阵

$$R^4=R^2\bigcirc R^2=\begin{bmatrix} 1 & & & & & & \\ 0.96 & 1 & & & & & \\ 0.98 & 0.96 & 1 & & & & \\ 0.99 & 0.96 & 0.98 & 1 & & & \\ 0.98 & 0.96 & 0.98 & 0.98 & 1 & & \\ 0.94 & 0.94 & 0.94 & 0.94 & 0.94 & 1 & \\ 0.95 & 0.95 & 0.95 & 0.95 & 0.95 & 0.94 & 1 \end{bmatrix} \quad (4)$$

可将其中的元素归并分类，当 λ 从 1 降为 0，由于 R_λ 不断变化，分类由细变粗逐渐归并，形成动态分类。

当取不同的置信水平时，各水体单元得

到分类。其依据是各水体单元受污染程度的相似性及差异性。当置信水平越高，水体单元得到越详细的分类；当置信水平越低，水体单元分类的模糊性越大。当取λ=0.98、0.96时武水河宜章江段7个水体单元的分类如下：

$[A_1]$，$[A_2]$，$[A_3、A_4、A_5]$，$[A_6]$，$[A_7]$；

$[A_1、A_2]$，$[A_3、A_4、A_5]$，$[A_6]$，$[A_7]$。

武水河宜章江段水体单元动态分类聚类如图1所示。同时，可根据相似矩阵中每行最大值即置信水平，将水体单元分级（表1）

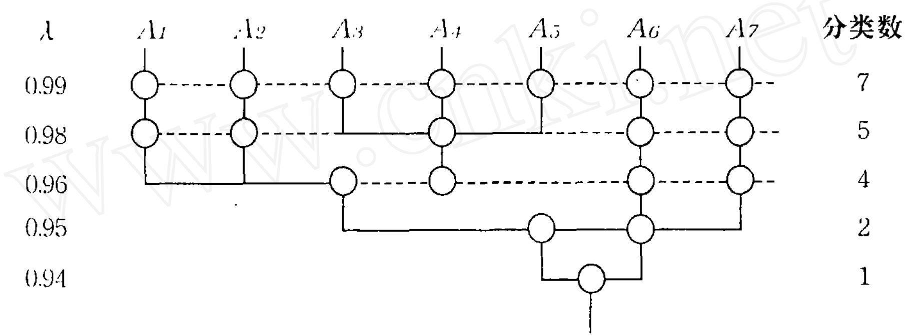


图1　武水河宜章江段水体单元动态分类聚类

表1　武水河宜章江段水体单元的分级

序号	1	2	3	4	5	6	7
水体单元	A_1	A_4	A_5	A_3	A_2	A_7	A_6
λ_i	0.99	0.99	0.98	0.90	0.96	0.95	0.94
分类	Ⅰ	Ⅰ	Ⅱ	Ⅱ	Ⅲ	Ⅳ	Ⅳ

根据以上分析看出，武水河宜章江段7个断面其污染程度的排序为$A_6>A_7>A_2>A_3$、$A_5>A_1$、A_4。第1第4两个断面污染程度相近，属于受污染最轻的断面；第6断面位于两个氮肥厂排污口下游300 m左右，因此是污染最严重的一个断面；第7断面虽距第6断面有较长的距离，得到一定自净，但由于期间又有偏溪河这一较大排污支流的汇入，使这种自净作用被抵消一部分。因此，其污染程度仅次于第6断面；污染程度位居第3的为第2断面，这是因为第2断面上游200 m即为杨家河排污支流，杨家河是水量较大，污染较重的排污口，因此使第2断面污染相应加重。第2断面距第4断面有一段较长的距离。其间虽也有排污，由于其水量和污染物浓度不大，因此自净作用超过排污作用，使水质有明显好转。

采用式（1）所列最大最小法公式计算模糊相似关系（表2）。

将表2的模糊相似短阵转换为模糊等价矩阵（式（5））。

当取不同的置信水平λ时，各污染物依据其污染程度的大小得到聚类。

当λ从1取到0，污染物的污染程度得到由细到粗的分类，其分类表现为一个动态过程，各污染物的聚类如下：

λ=1，根据污染程度，污染物被聚为7类，即[DO、BOD_5、COD、NO_2-N、NO_3-N、CN^-、Pb、Zn]；[SS]；[NH_3-N]；[Φ-OH]；[As]；[F^-]；[细菌总数]。

λ=0.90，污染物被聚为3类，即[DO、BOD_5、COD、NO_2-N、Φ-OH、CN^-、As、Pb、Zn、SS、F^-]；[NH_3-N]；[细菌总数]。

表2　污染因子的模糊相似矩阵

	污染因子	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	DO	1													
2	BOD_5	1	1												
3	COD	1	1	1											
4	SS	0.93	0.93	0.93	1										
5	NH_3-N	0.04	0.04	0.04	0.04	1									
6	NO_2-N	1	1	1	0.93	0.04	1								
7	NO_3-N	1	1	1	0.93	0.01	1	1							
8	Φ−OH	0.94	0.94	0.94	0.89	0.04	0.91	0.94	1						
9	CN^-	1	1	1	0.93	0.04	1	1	0.94	1					
10	F^-	0.87	0.87	0.87	0.90	0.04	0.87	0.87	0.85	0.87	1				
11	As	0.94	0.94	0.94	0.90	0.04	0.94	0.94	0.90	0.94	0.89	1			
12	Pb	1	1	1	0.93	0.04	1	1	0.94	1	0.87	0.94	1		
13	Zn	1	1	1	0.93	0.04	1	1	0.94	1	0.87	0.94	1	1	
14	细菌总数	0.26	0.06	0.26	0.28	0.15	0.26	0.26	0.26	0.27	0.30	0.27	0.26	0.26	1

$$R^4=\begin{bmatrix}
1 & 1 & 1 & 0.93 & 0.15 & 1 & 1 & 0.94 & 1 & 0.90 & 0.94 & 1 & 1 & 0.30\\
1 & 1 & 1 & 0.93 & 0.15 & 1 & 1 & 0.94 & 1 & 0.90 & 0.94 & 1 & 1 & 0.30\\
1 & 1 & 1 & 0.93 & 0.15 & 1 & 1 & 0.94 & 1 & 0.90 & 0.94 & 1 & 1 & 0.30\\
0.93 & 0.93 & 0.93 & 1 & 0.15 & 0.93 & 0.93 & 0.93 & 0.93 & 0.90 & 0.93 & 0.93 & 0.93 & 0.30\\
0.15 & 0.15 & 0.15 & 0.15 & 1 & 0.15 & 0.15 & 0.15 & 0.15 & 0.15 & 0.15 & 0.15 & 0.15 & 0.15\\
1 & 1 & 1 & 0.93 & 0.15 & 1 & 1 & 0.94 & 1 & 0.90 & 0.94 & 1 & 1 & 0.30\\
1 & 1 & 1 & 0.93 & 0.15 & 1 & 1 & 0.94 & 1 & 0.90 & 0.94 & 1 & 1 & 0.30\\
0.94 & 0.94 & 0.94 & 0.93 & 0.15 & 0.94 & 0.94 & 1 & 0.94 & 0.90 & 0.94 & 0.94 & 0.94 & 0.30\\
1 & 1 & 1 & 0.93 & 0.15 & 1 & 1 & 0.94 & 1 & 0.90 & 0.94 & 1 & 1 & 0.30\\
0.90 & 0.90 & 0.90 & 0.93 & 0.15 & 0.90 & 0.90 & 0.90 & 0.90 & 1 & 0.90 & 0.90 & 0.90 & 0.30\\
0.94 & 0.94 & 0.94 & 0.93 & 0.15 & 0.94 & 0.94 & 0.94 & 0.94 & 0.90 & 1 & 0.94 & 0.94 & 0.30\\
1 & 1 & 1 & 0.93 & 0.15 & 1 & 1 & 0.94 & 1 & 0.90 & 0.94 & 1 & 1 & 0.30\\
1 & 1 & 1 & 0.93 & 0.15 & 1 & 1 & 0.94 & 1 & 0.90 & 0.94 & 1 & 1 & 0.30\\
0.30 & 0.30 & 0.30 & 0.30 & 0.15 & 0.30 & 0.30 & 0.30 & 0.30 & 0.30 & 0.30 & 0.30 & 0.30 & 1
\end{bmatrix} \quad (5)$$

根据以上分析可以看出，污染物的聚类，不仅反映了污染物间的相互依存关系，即某一污染物程度取决于或依赖于其他一种或几种污染因子，而且也反映了哪一类污染源对水体的作用最大。

2.2　污染参数的模糊相关分析

根据式(2)计算各污染因子的模糊相关矩阵(表3)。由表3看出，DO、BOD_5、COD、NO_2-N、NO_3-N、CN^-、Pb、Zn、Φ−OH、As、F^-之间相关系数在0.98以上，NH_3-N

表3 污染因子的相关性

污染因子	DO	BOD_5	COD	SS	NH_3 −N	NO_2 −N	NO_3 −N	Φ−OH	CN^-	F^-	As	Pb	Zn	细菌总数
DO	1													
BOD_5	1	1												
COD	1	1												
SS	0.98	0.98												
NH_3-N	0.33	0.33	0.33	0.39	1									
NO_2-N	1	1	1	0.98	0.33	1								
NO_3-N	1	1	1	0.98	0.33	1	1							
Φ−OH	0.98	0.98	0.98	0.96	0.38	0.98	0.98	1						
CN^-	1	1	1	0.98	0.33	1	1	0.98	1					
F^-	0.97	0.97	0.97	0.99	0.43	0.97	0.97	0.97	0.97	1				
As	0.99	0.99	0.99	0.97	0.38	0.99	0.99	0.97	0.99	0.98	1			
Pb	1	1	1	0.98	0.33	1	1	0.98	1	0.97	0.99	1		
Zn	1	1	1	0.98	0.33	1	1	0.98	1	0.97	0.99	1	1	
细菌总数	0.52	0.52	0.52	0.58	0.94	0.52	0.52	0.57	0.52	0.64	0.57	0.52	0.52	1

与其它污染物相关性较弱，但与细菌总数的相关性较强。以上污染物的相关性，既反映了各类污染物的来源，又反映了各污染因子在水体中的相互依赖程度。

3 结果与讨论

3.1 通过用最大矩阵元法对水体单元和污染因子进行聚类分析，同时结合相关分析对污染因子间的关系进行初步探讨，可以看出：由于各类污染物的相互依赖关系不同及其来源不同，导致污染程度的不同，可以通过模糊聚类分析这一数学过程得到体现。

3.2 从对污染因子聚类的结果看，对水体中的污染物不能孤立地看。由于每个污染因子都具有一定的来源，同时污染因子彼此间存在物理化学作用而建立起的相互依赖关系，使得水体中污染参数具有类别的优势和弱势。就武水而言，NH_4-N 和细菌总数占优势。

3.3 通过污染物的聚类分析，不仅揭示了污染物的相互依存关系，还揭示了哪一类污染源对纳污水体的影响最大，即间接评价了各类污染源对环境的污染程度，将水质与污染源的关系直观地表征了出来。

3.4 从各污染因子的相关矩阵看出，各污染因子之间的相互关系的网络结构和层次结构，这对于寻找治理污染的途径是很有用的。如降低悬浮物的输入，有可能降低氟化物及 NO_2-N、NO_3-N 的输入，从而降低了 BOD_5、COD，增加了DO，减少了 NH_3-N 的污染。

3.5 从污染因子的相关角度看，武水宜章江段SS与 NO_2-N、NO_3-N、F^-、Pb、Zn等的相关系数在0.98以上，也许可表明水中的 NO_2-N、NO_3-N、F^-、Pb、Zn等污染物的来源与面污染源有密切关系。该区有部分区域是高氟区，区域内有较多小型黑色金属与有色金属矿点，这种工业特点也进一步表明了这一点。

参 考 文 献（略）

收稿日期：2000—05—12